Annals of Agrarian Science

Journal homepage: http://journals.org.ge/index.php

Verification of WRF simulated accumulated precipitation

N. Kutaladze[a]*, G. Mikuchadze[a], M. Tatishvili[b]

[a]Ministry of Environmental Protection and Agriculture of Georgia, National Environmental Agency; 150, David Agmashenebeli Ave., Tbilisi, 0112, Georgia

[b] Georgian Technical University, Institute of Hydrometeorology; 150-a, David Agmashenebeli Ave., Tbilisi, 0112, Georgia

Received: 05 October 2020; accepted: 12 October 2020.

A B S T R A C T

Purpose of this study is investigation of Advanced Research Weather Forecasting Model's (WRF-ARW) skill in Quantitative Precipitation Forecasting for Georgia's conditions, where orographic features play key role in modeling convectional processes. The Country territory is prone to flash floods and mudflows, Quantitative Precipitation Estimation (QPE) and Quantitative Precipitation Forecast (QPF) on any leading time are very important for Georgia. We have analyzed several convection parameterization and microphysical schemes combination for different rainy episodes and heavy rainy phenomena. We estimate errors and biases in accumulated 6 h precipitation using different spatial resolution during model performance verification for 12-hour and 24-hour lead time against corresponding rain gouge observations and satellite data. Various statistical parameters have been calculated for the 8-month comparison period and some skills of model simulation have been evaluated.

Keywords: Air mass convection, Numerical weather prediction (NWP), Quantitative Precipitation Forecasting (QPF), False alarm, Extremal Dependence Index, Contingency table.

Acknowledgements: The research was funded by the Shota Rustaveli National Science Foundation (SRNSF) of Georgia (grant FR17_529).

*Corresponding author: Nato Kutaladze: E-mail address: cwlam08@gmail.com

Introduction

Precipitation forecasts are one of the most demanding applications in numerical weather prediction (NWP). Georgia, as the whole Caucasian region is characterized by very complex topography. Such complex character of the relief and the vicinity of the Black and Caspian seas considerably deforms large-scale (synoptic) processes and causes formation of local and regional peculiarities of atmospheric processes and strong spatial inhomogeneity of meteorological fields and mainly determines the precipitation regime all over the territory of Georgia. The climatic picture totally differs in both parts of Georgia as divided by the Likhi Range. The Black Sea influences the climate of West Georgia, resulting in abundant precipitation. The climate in the plains of East Georgia is dry. The annual amounts of precipitation vary in the range of 400-600 mm in the plains, and 800-1,200 mm in the mountains.

Taking into account above mentioned orographic and synoptic complexity simulation results from regional NWP do not have the same quality for all areas within the domain. This is especially evident for precipitation quantitative forecasts, where in contrast to other forecast variables, generally no significant improvement could be achieved in the last decade [1-4]. Quantitative Precipitation Forecasting (QPF) on regional scales is still inadequate for many applications such as in hydrology and flood forecasting. For this purpose, it is essential to simulate precipitation accurately down to the size of small catchment areas. Our focus is on the formation and organization of convective precipitation systems in a low-mountain region. Several problems in connection with QPF have been identified for mountain regions, which include the overestimation and underestimation of precipitation on the windward and lee side of the mountains, respectively, and a phase error in the diurnal cycle of precipitation leading to the onset of convective precipitation in model forecasts several hours too early [5-7].

Many countries have national rain gauge networks that provide observations that can be used to verify the model QPFs. Since rainfall depends strongly on atmospheric motion, moisture content, and physical processes, the quality of a model's rainfall prediction is often used as an indicator of overall model health [8-10].

Purpose of this study is investigation of Advanced Research Weather Forecasting Model's (WRF-ARW) skill in Quantitative Precipitation Forecasting for Georgia's conditions, where orographic features play key role in modeling convectional processes, testing several convection parameterization and microphysical schemes, as well as estimation errors and biases in accumulated 6 h precipitation using different spatial resolution for verification model performance against observations.

1. Data and method

The strategy for any forecast verification application includes certain rational steps: choosing and matching a set of forecast/observation pairs, defining the technique to compare them, aggregating (pooling) and/or stratifying the forecast/observation pairs in appropriate data samples, applying the relevant verification statistics and, ultimately, interpreting the scores, not forgetting to analyze the statistical significance of the gained results. Deterministic QPFs can be formulated and taken as either categorical events or continuous variables and verified correspondingly utilizing respective verification approaches and measures [11-13]. Verifying QPFs as categorical events is clearly more common. The categorical approach involves issues like whether or not it rained during a given time period (rather than at a given instant) or, alternatively, whether the rainfall amount exceeded a given threshold. Verifying rainfall amount as a continuous variable brings about certain caveats because the rainfall amount is not a normally distributed quantity. Very large rainfall amounts may be produced by a forecasting system and, then again, in some cases very little or no rain. Many of the verification scores for continuous variables, especially those involving squared errors, are very sensitive to large errors. Consequently, categorical verification scores provide generally more meaningful information of the quality of the forecasting systems (or skill of the human forecasters) producing QPFs [14-15].

The WRF-ARW version 3.1 model was running operationally on the NEA cluster during several months with two different model configuration combining different convective and microphysical schemes. The simulation was run using the Betts-Miller-Janjic (BMJ) parameterization, which is based on the Betts-Miller convective adjustment scheme, Betts and Miller (1986), with Lin microphysics scheme, for mother domain and the same convection scheme with WRF single-moment 6 class microphysics scheme for nest. All simulations used the Yonsei University Planetary boundary layer schemes, 5-layer soil model for Surface layer and Dudhia's Shortwave and RRTM Long wave radiation

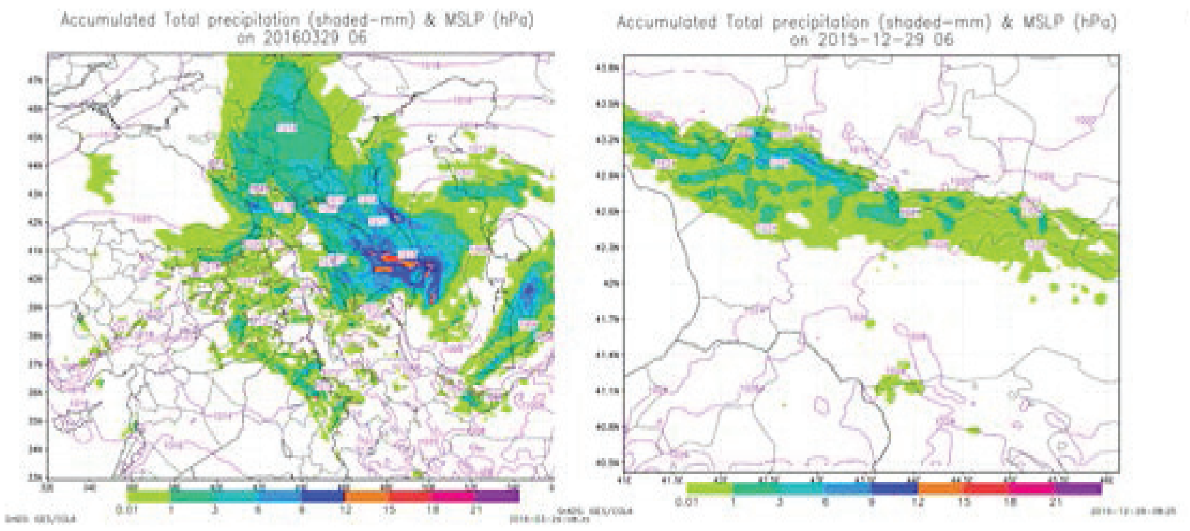

Fig. 1. *6 h total accumulated precipitation from mother and nested domain of WRF-ARW model.*

As station data are point based and model output data are gridded we extracted time series from model output GRIB files for comparison schemes. The model performance was carried out with two way nesting option.

For the study 6 hourly accumulated precipitation sums from 70 automated weather stations (AWS)/ posts and rain gouges were used (see fig.1). The number and spatial coverage of precipitation measurement points is not sufficient for the characterizing precipitation pattern across the country. Especially, when in mountainous parts of country we have very few stations or none at all.

For cross validation of measured precipitation satellite TRMM (The Tropical Rainfall Measuring Mission) data files (are nearly real time data in binary format) from the NASA web site (ftp://trmmopen.gsfc.nasa.gov/pub/merged/3B42RT) was investigated. TRMM mainly observe rain structure, rate and distribution in tropical and subtropical region, the data play an important roll for understanding mechanisms of global climate change and monitoring environmental variation. TRMM is the first space mission dedicated to measuring tropical and subtropical rainfall through microwave and visible/infrared sensors, including the first space borne rain radar [16].

2. Verification metrics and Calculation

The WRF–ARW model, version 3.1, was used for simulation. The integration domain covered a roughly 1500 km x1500 km region centered over the south-Caucasus region. Mother Domain overlaps S 300- N 500 and E 300- W 600 territory, with resolution 9 km and two-way nest focused mostly on Georgia and its resolution is 3 km. Forecast was integrated at 0000 UTC 1 April 2015, this month was too rainy this year (25 days >1mm) and synoptic processes developed here varied in wide range. In the present study all runs were initialized with 25-km NCEP GFS Model GRIB data, and integrated 72 h. forecast for each day.

Comparisons are made with the surface observation data 12 h accumulated total precipitation which unfortunately very pars for this aim. Distribution of stations are given on the picture 1. For each case errors and deviations were estimated for individual observation point on the all model simulated forecast valid time: 15h; 27h; 39h; 51h; 63h; (corresponding to local observation), for any configuration and both domains.

There are a number of ways to quantify the verification of NWP models. In this paper, categorical as well as continuous statistics have been applied. Categorical statistics can be derived from two-dimensional contingency tables. We determine the accuracy of the model to distinguish between rain and no rain by using contingency tables (Table 1).

This two dimensional table that gives the discrete joint sample distribution of forecasts and observations in terms of cell counts. Cell count h is the number of event forecasts that correspond to event observations or the number of hits, cell count f is the number of event forecasts that do not correspond to observed events or the number of false alarms, cell count

m is the number of no-event forecasts corresponding to observed events or the number of misses, and cell count z is the number of no-event forecasts corresponding to no events observed or the number of zeros. The forecast quality for this (2×2) binary situation can be assessed using a large number of different verification measures [17-19].

HR is a categorical forecast score equal to the total number of correct event forecasts (hits) divided by the total number of estimates:

HR = correct estimates/total estimates = $(h + z)/n$ (1)

POD is the fraction of those occasions where the estimation event really occurred:

POD = correct rain estimates/rain observation $=h/(h + m)$ (2)

It ranges from zero to one, where one indicates a perfect forecast. POD is very sensitive to the climatological frequency of the event. It can be

Table 1. *Two dimensional contingency table*

Forecast/observed	Yes	No	Sum
Yes	Hit (h)	False alarm (f)	Forc. yes (h + f)
No	Miss (m)	Correct non-event (z)	Forc. no (m + z)
Sum	Obs. Yes (h+m)	Obs. No (f+z)	(n=h+f+m+z)

improved by issuing more “yes” forecasts to increase the number of hits. It also ignores FAR.

FAR is equal to the number of false alarms divided by the total number of event forecast:

FAR = false rain estimates/rain estimates = f /(h + f) (3)

It varies from one to zero, where zero indicates a perfect forecast. It is sensitive to false alarms, but ignores misses. It is also very sensitive to the climatological frequency of the event.

Bias score is equal to rain estimates divided by the rain observations. It can be positive and negative and ranges from zero to infinity. It is an objective measure, which mainly addresses precipitation spatial distribution. A positive bias score shows that the model overestimates the observation values while negative bias score indicates an underestimation over the investigated area.

Bias = rain estimates/rain observations = (h + f)/(h + m) (4)

Skill score is the relative accuracy of the forecast over some reference forecast. The reference forecast is generally an unskilled forecast such as random chance, persistence defined as the most recent set of observations which implies no change in condition, or climatology. Skill score refers to the increase in accuracy purely due to the “smarts” of the forecast system. There are as many skill scores as there are possible scores and they are usually based on the expression:

$$SS = (A_{estimated} - A_{reference})/(A_{perfect} - A_{reference}) \quad (5)$$

The skill scores generally lie in the range 0 to 1, where the skill score of 1 indicates a forecast better than reference forecast and 0 indicates it is not better than the reference forecast. In this study, we have used the TSS as skill score.

It is computed from:

TSS = (hz – fm)/((h+m)(f+z)) (6)

It ranges from −1 to 1, where 1 indicates a perfect forecast. This kind of skill score has been used in this study because it is independent of dry or wet regime (Ebert and McBride 1997). Other skill scores, like the Heidke skill score, depend on the current precipitation distribution. In this case, the reference forecast is based on random chance [20-21].

In order to quantify the verification results, continuous statistics have been used. Continuous verification measures comprise mean error (ME), MAE, and root mean square error (RMSE). By using multicategory contingency tables with three different thresholds (>5 mm; from 5 to 20 mm and <20 mm) the accuracy of prediction for each threshold has been determined.

The metrics based on continues statistics such as correlation and RMSE were calculated also for GFS global model. Precipitation simulated with WRF-ARW model mother and nested domains was verified using with all statistical parameters described above.

3. Results and discussions

For each station, observed and model simulated time series were compared and the mentioned statistical parameters were calculated.

On fig, 3a, verification results of global model GFS versus observation from several stations are presented. This metrics varies from station to station, but both of them are not high correlations are below 0.7 and standardized deviations values also near to 1. So the GFS model scale for simulated precipitation cannot satisfy local prediction needs.

a.) b.) c,)

Fig. 2. *Tay lor Diagram for some stations and GFS (a); ME, and MAE of WRF-ARW precipitation forecast*

over 5 subregions in Georgia (b) as well as POD and FAR (for rain/no rain)(c)

Analyzing the obtained results we grouped the stations according to the stations elevation and synoptic processes in the country. Number of subregions in which the metrics have approximately the same range is five: South east of Great Caucasus (SEC), South West Great Caucasus (SWC), Plain of Eastern Georgia (EP), Plain of Western Georgia (WP) and Southern Georgia (SG). This statistical parameters varies from month to month and very defending to thresholds, also. Most of the synoptic processes that take place in the country during the year are caused by the impact of incoming air mases from the Black Sea, which is often accompanied by precipitation in western Georgia.

From the Fig 3 b.) & c.) ME, MAE, POD and FAR indicates better performance in the subdomain WP, which was to be expected. This is also facilitated by the fact that 80% of the daily precipitation in this area is in the range of 5 to 20 mm, and simulating such an amount of precipitation by the model is particularly good, also most of the stations are mainly located in lowland areas, reducing observation distribution errors. (*Table 2.)*

As it is shown at the table 2 model performance metrics are better for the 12 hour lead time data, for the station as well as satellite observations. In fact, it was expected. Results are better when comparing model to TRMM than model versus stations. It should be mentioned, that while comparing observed data from satellite to station, deviations are even more than deviation from the forecasted precipitation and station observations. Obtained results might be explained by the different nature of precipitation data, gridded ones from model and satellites and point based observations from surface stations, with quite pares coverage of territory.

4. Conclusion

The results of model validation are not homogeneous inside of domain and are highly dependent on the physical content of the synoptic process. The results are characterized by a certain seasonality. We compare simulation results of some rainy event with several physical schemes configuration and different combination for particular case improves the results, although using this configuration for other processes gives worse results. Which emphasizes that in the future it is desirable to simulate not one configuration of the model, but several simultaneously and generate a multi model ensemble, which also need very long-term and careful validation and verification with additional metrics. In addition, it is necessary to systematically verify all available observational information and allow much longer periods for comparison.

Table 2. *Results of comparison of 12 h and 24 h lead time model forecasted precipitation and stations measurement and TRMM*

	Mode vs stations (12 h)	Mode vs stations (24 h)	Mode vs. TRMM (12h)	Mode vs. TRMM (12h)
POD	0.781983	0.703396	0.82967	0.761895
Bias	0.278841	0.327582	0.238917	0.266863

References

[1] D.-Q. Nguyen, J. Renwick, and J. Mcgregor, Variations of surface temperature and rainfall in Vietnam from 1971 to 2010, International Journal of Climatology, vol. 34 no.1 (2014) 249-264.

[2] I. Jankov;, J. Berner, J. Beck, H. Jiang, J. B. Olson, G. Grell, T. G. Smirnova, S. G. Benjamin, J. M. Brown, A Performance Comparison between Multiphysics and Stochastic Approaches within a North American RAP Ensemble. Mon. Wea. Rev. (2017) 145 (4): 1161–1179. https://doi.org/10.1175/MWR-D-16-0160.1

[3] N. Rahmawati & Maciek W. Lubczynski, Validation of satellite daily rainfall estimates in complex terrain of Bali Island, Indonesia, Theoretical and Applied Climatology, v.134 (2018) 513–532.

[4] G. A. Grell and D. Devenyi, A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. Geophys. Res. Lett., 29 (2002) 1693-1696.

[5] J. S. Kain, The Kain-Fritsch convective parameterization: An update. J. Appl. Meteor., 43, (2004)170-181.

[6] J. S. Kain and J. M. Fritsch, A one dimensional entraining/detraining plume model and its application in convective parameterization, J. Atmos. Sci. 47 (1990) 2784-2802.

[7] J. S. Kain, and J. M. Fritsch, Convective parameterization for mesoscale models: The Kain-Fritsch scheme, The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr., Amer. Meteor. Soc., No 24 (1993) 165-170.

[8] J. Dudhia, Numerical study of convection observed during winter monsoon experiment using a mesoscale two dimensional model., J. Atmos. Sci. 46 (1989) 3077–3107.

[9] A. K. Betts, and M. J. Miller, The Betts-Miller scheme, The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr., Amer. Meteor. Soc., No. 24 (1993) 107-121.

[10] Z. I. Janjic, The step-mountain eta coordinate model: Further development of the convection, viscous sublayer, and turbulence closure schemes, Mon. Wea. Rev., 122 (1994) 927-945.

[11] J. Dudhia, A multi-layer soil temperature model for MM5. in: Preprints, The Sixth PSU/NCAR Mesoscale Model Users' Workshop, Boulder/CO, Boulder/CO. NCAR (1996).

[12] E.E. Ebert, U. Damrath, W.Wergen, M.E. Baldwin, The WGNE assessment of short-term quantitative precipitation forecasts, Bull. Amer. Meteor. Soc., 84 (2003) 481–492.

[13] A. Gevorgyan, Verification of daily precipitation amount forecasts in Armenia by ERA-Interim model. International Journal of climatology V.33 I.12 (2013) 2706-2712.

[14] E.E. Ebert and W.A. Gallus, Toward better understanding of the contiguous rain area (CRA) method for spatial forecast verification. Wea. Forecasting, 24, (2009)1401-1415.

[15] A. M. Noorian, S. Sodoudi, G. Manfred, E. Reimer, Daily precipitation forecast of ECMWF verified over Iran. Theoretical and Applied Climatology 99(1) (2010) 39-51 DOI: 10.1007/s00704-009-0118-9

[16] D. D. Tien, L. R. Hole, D. T. Anh, C. H. Duc, Th. N. Ba, «Verification of Forecast Weather Surface Variables over Vietnam Using the National Numerical Weather Prediction System», Advances in Meteorology, vol. 2016, Article ID 8152413 (2016) https://doi.org/10.1155/2016/8152413

[17] X. Guoqiang, L. Xudong, Y. Hui, H. Liping, X. Jishan, Precipitation simulation using different cloud-precipitation schemes for a landfall typhoon. Plateau Meteorol., 26 (2007) 891-900.

[18] A. Arakawa, The cumulus parameterization problem: Past, present, and future, J. Climate, 17 (2004) 2493-2525.

[19] T. Ngo-Duc, C. Kieu, M. Thatcher, D. Nguyen-Le, and T. Phan-Van, Climate projections for Vietnam based on regional climate models, Climate Research, vol. 60, no.3 (2014) 199–213.

[20] C.A.T. Ferro and D.B. Stephenson, Extremal Dependence Indices: improved verification measures for deterministic forecasts of rare binary events, Wea, Forecasting 26 (2011) 699-713.

[21] C. Pennelly, G. W. Reuter, T. Flesch, Verification of the WRF model for simulating heavy precipitation in Alberta, J. Atmospheric Research s 135–136 (2014) 172–192 DOI: 10.1016/j.atmosres.2013.09.004